

Performance Metrics, KPI Distortion, and Campbell's Law

Why Indicators Stop Reflecting the Realities They Were Built to Measure

AI Systems and Reality Drift - Note #6 v1

A. Jacobs - Semantic Fidelity Project

The Basic Pattern

Modern organizations depend on measurement because they cannot observe every outcome directly. At institutional scale, social outcomes, school performance, company operations, and healthcare quality have to be translated into indicators before they can be managed. Statistics, testing systems, dashboards, KPIs, and performance metrics become the organization's way of seeing.

Those indicators are useful, but they are also substitutions. A graduation rate may stand in for educational quality. A response time may stand in for service quality. A KPI may stand in for organizational effectiveness. Each one makes a complex reality easier to track, but also easier to mistake for the reality itself.

At first, these indicators provide visibility, improve coordination, and help organizations make decisions. Over time, however, a familiar pattern can emerge. As funding, promotions, reputation, and evaluation become tied to the indicator, people begin optimizing the measurement itself. The representation remains visible, but the underlying reality becomes harder to see.

When Measurement Becomes the Target

Indicators are designed to represent conditions that are difficult to observe directly. The measurement is not the outcome. It is a representation of the outcome.

At first, the relationship may remain reasonably stable. A test score can still reflect learning, a productivity metric can still reflect real output, and a performance indicator can still point toward better results. The measurement is imperfect, but it remains close enough to the reality it is meant to describe.

But once an indicator becomes important enough, behavior starts to reorganize around it. People learn what the measurement rewards. Organizations adjust to what the evaluation notices. The metric stops being only a way to observe the system and becomes something the system is trying to satisfy.

This is when measurement becomes the target. Improvement in the indicator no longer guarantees improvement in the underlying reality. The metric may continue rising while the outcome becomes less certain.

How Indicator Drift Emerges

Stage 1 — Observation

An organization introduces a measurement system. At this stage, the indicator remains reasonably connected to the underlying condition it is meant to represent.

Stage 2 — Incentivization

The measurement becomes tied to funding, rewards, accountability, status, or performance evaluation. Improving the metric becomes increasingly valuable.

Stage 3 — Optimization

Individuals and organizations adapt their behavior to improve the indicator. Effort increasingly focuses on the measurement itself, and strategies emerge that improve the metric more efficiently than improving the underlying outcome.

Stage 4 — Drift

The measurement continues improving, reports remain positive, and the dashboard continues showing success. Yet the relationship between the indicator and the reality it represents begins weakening. The representation appears healthy while the reality becomes harder to assess.

The absence of failure is not proof of alignment. It may only mean the measurement system has stopped detecting the gap.

Semantic Fidelity and Constraint Collapse

At this point, the measurement system has not necessarily failed on its own terms. What weakens is the structural relationship between the indicator and the reality being measured.

A metric may preserve the surface appearance of measurement while losing the conditions, relationships, and constraints that made the measurement meaningful. This is where indicator drift becomes a semantic fidelity problem. The representation remains coherent, but its connection to the underlying reality has degraded.

Indicator drift also involves constraint collapse. Measurement is supposed to constrain judgment by keeping decisions tied to observable outcomes. But when organizations become organized around the metric itself, that constraint can weaken. The system remains measured, but measurement loses some of its corrective force.

The Representation Stack

Indicator drift becomes more serious when measurements move through a wider representation stack. Reality becomes a metric, the metric becomes a dashboard, and the dashboard becomes evidence for broader claims about performance or success. Each layer makes the system easier to manage while increasing the distance from the original condition.

This is how representational failure can become organizational. A metric may begin as limited evidence about reality, but after it moves through reports, incentives, decisions, and institutional trust, it can harden into a belief about the world. The further the representation travels from the reality it was meant to describe, the easier it becomes for measurement to be mistaken for outcome.

Examples Across Systems

Education: Standardized tests may be introduced to measure learning, but as accountability increases, schools can become more focused on improving test performance than improving learning itself.

Corporate Performance Metrics: Organizations use KPIs to monitor productivity and effectiveness, but employees and departments may learn to improve the indicator without improving the underlying work.

Healthcare Systems: Quality indicators can help healthcare organizations monitor performance, but attention may shift toward documented performance while less measurable aspects of patient care become harder to see.

Public Policy: Governments use statistical indicators to evaluate programs and institutions, but once those indicators shape funding and accountability, organizations may optimize the measurement more than the underlying problem.

Conclusion

Campbell's Law is not only a problem of bad metrics. It is a representational problem. Metrics, indicators, dashboards, KPIs, rankings, and evaluation systems make large-scale coordination possible by translating complex realities into manageable forms.

The deeper failure occurs when the measurement remains visible, trusted, and optimized while its relationship to the underlying reality weakens. The challenge is maintaining fidelity between the indicator and the reality it was designed to represent. A system remains aligned only when its measurements stay answerable to the outcomes they were created to track.

Related Phrases and Concepts

- Campbell's Law
- Goodhart's Law

- KPI distortion
- metric gaming
- target corruption
- performance indicators
- accountability metrics
- indicator corruption
- teaching to the test
- metric substitution

Semantic Fidelity Project Resources

- [Research Library \(GitHub\)](#)
- [Articles & Essays \(Substack\)](#)
- [Semantic Fidelity Glossary](#)
- [LLM Failure Modes and Semantic Fidelity](#)